



HAL
open science

Aggregate production planning in the automotive industry with special consideration of workforce flexibility

Thomas Sillekens, Achim Koberstein, Leena Suhl

► **To cite this version:**

Thomas Sillekens, Achim Koberstein, Leena Suhl. Aggregate production planning in the automotive industry with special consideration of workforce flexibility. *International Journal of Production Research*, 2010, pp.1. 10.1080/00207543.2010.524261 . hal-00647964

HAL Id: hal-00647964

<https://hal.science/hal-00647964>

Submitted on 4 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Aggregate production planning in the automotive industry
with special consideration of workforce flexibility**

Journal:	<i>International Journal of Production Research</i>
Manuscript ID:	TPRS-2010-IJPR-0358.R1
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	04-Aug-2010
Complete List of Authors:	Sillekens, Thomas; Daimler AG Koberstein, Achim; University of Paderborn, Decision Support and Operations Research Lab Suhl, Leena; University of Paderborn, Decision Support and Operations Research Lab
Keywords:	AGGREGATE PLANNING, MIXED INTEGER LINEAR PROGRAMMING, FLEXIBILITY
Keywords (user):	Automotive Industry

SCHOLARONE™
Manuscripts

International Journal of Production Research
Vol. 00, No. 00, 00 Month 2009, 1–24

ARTICLE

Aggregate production planning in the automotive industry with special consideration of workforce flexibility

Thomas Sillekens^a, Achim Koberstein^{b*} and Leena Suhl^b

^a*Daimler AG - Group Research and Advanced Engineering, Ulm, Germany;*

^b*Decision Support and Operations Research Lab (DSOR-Lab) - University of
Paderborn, Germany*

(Received 00 Month 200x; final version received 00 Month 200x)

We present a new mixed integer linear programming approach for the problem of aggregate production planning of flowshop production lines in the automotive industry. Our model integrates production capacity planning and workforce flexibility planning. In contrast to traditional approaches it considers discrete capacity adaptations which originate from technical characteristics of assembly lines as well as from work regulations and shift planning. In particular, our approach takes change costs into account and explicitly represents a working time account via a linear approximation. A solution framework containing different primal heuristics and preprocessing techniques is embedded into a decision support system. Finally, we present an illustrative case study and computational results on problem instances of practically relevant complexity.

Keywords: Automotive Industry, Aggregate Production Planning, Mixed Integer Programming, Workforce Flexibility

1. Introduction

Today, the automotive industry is challenged by high competition, high product variety and shortened product life-cycles. Consequences are high cost pressure and a highly dynamic product demand. One way of reacting to these challenges is improving the utilization of flexibility inherent in a given production network. In this paper, we consider the planning problem of operating installed flowshop production lines efficiently at optimal costs for a midterm planning horizon of typically 12 to 24 month. In the automotive industry material costs are the biggest cost factor which account for about half of the

*Corresponding author. Email: koberstein@dsor.de

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

cost of sales. However, they cannot be influenced by production planning. They are followed by labor costs ranging from 15 to 20% of the sales. Furthermore, some authors suggest that up to 50% of all occurring costs along the entire supply chain are labor costs (see Becker 2006). To be able to utilize labor capacity efficiently, there is a trend that workforce and working time hours are also adaptable to become more flexible. Examples are working time accounts, varying hours in shift work and temporary workers. This flexibility in workforce and working time is used very widely in the entire production industry throughout Europe today (see Bauer *et al.* 2007). However, regulations regarding all these measures make the resulting planning task even more challenging. Due to these reasons it is important to plan production and workforce capacity in an integrated way. Moreover, stock and buffers within the production structure play an important role for the problem. They have to be considered to ensure feasible production plans. All of these requirements put the problem for the mid term planning horizon in the field of aggregate production planning (APP).

Since pioneers such as Holt *et al.* (1955), Holt *et al.* (1956) and Hanssmann and Hess (1960) founded the field, there has been a lot of research on the topic of aggregate production planning. An overview of earlier developments is given in Nam and Logendran (1992). However, very few authors consider the specific challenges of aggregate planning in the automotive industry. An approach of Oliff *et al.* (1989) considers crew loaded environments aside from a specific industry. Different heuristics have been developed over time, such as the production switching heuristic (see Orr 1962, Nam and Logendran 1995, Cha and Hwang 1996) or the search decision rule (see Taubert 1968, Gilgeous 1989) with the goal to be adaptable to practical problems. However, the development of workforce flexibility and regulatory requirements is not considered in detail. Only single aspects such as annualized hour agreements are considered in some works, e.g. by Corominas *et al.* (2007). Working time models and resulting shift work have not been addressed properly yet. This may be due to the fact that this topic has become especially urgent during the last decade. Most publications deal with the extension of existing models regarding stochastic aspects. This includes stochastic linear programming (see Leung and Wu 2004) and possibilistic linear programming (see Wang and Lang 2005) as well as a combination of fuzzy random variables, simulation and genetic algorithms (see Ning *et al.* 2006). Different studies of Buxey (2003) and Buxey (2005) have shown that the current approaches to aggregate production planning are not put into practice. He argues that they are too complicated or do not reflect the actual problem description to be applicable. More recently Silva *et al.* (2006) proposed an APP model which includes three performance criteria and three groups of constraints related with workforce, production and inventory level. In addition they developed a computerized decision support system (DSS) that searches and provides “the best solution”, if exists, for this type of problems. In the spite of the complexity of the model formulation the DSS does not require the user to have a high knowledge level of mathematical aspects of the model and allows managers through a user-friendly and interactive process, to generate, evaluate and compare different solutions for the APP problem. Also quite recently, a dynamic programming approach has been developed by Askar and Zimmermann (2006) and Askar *et al.* (2007) to cope with the special problem which the automotive industry is facing. In fact, it can be seen as a predecessor of our approach, as it was developed with the same industrial partner. Following the lines of dynamic programming, in this approach the space of feasible solutions is represented by a space of discrete states, which are computed recursively to determine an optimal path through the state space. It was shown by the authors, that some instances of the planning problem can be solved efficiently by using this approach.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

However, the acceptance of the approach in practise suffered from some drawbacks, some of which are principle drawbacks of dynamic programming based solution approaches and some of which are induced by the special structure of the problem at hand:

- In the dynamic programming approach, some model objects which are of continuous nature or can have large integer values, have to be represented by a set of discrete states. This inevitably leads to a loss of precision. Especially buffers between production lines could hardly be represented with sufficient accuracy.
- The consideration of change costs led to a vastly extended state space.
- To maintain solvability within acceptable computing times, it was necessary to reduce the search space of the problem by state aggregation, which turns the method into a heuristic. At the same time, the approach does not provide lower bounds, so the quality of a solution is difficult to evaluate for the planner.
- In the dynamic programming approach, there is no separation between model formulation and solution algorithm. As a consequence, adopting the dynamic programming approach to slightly different planning situations turned out to be a very laborious and complex task.

Based on our literature review, our experience with the dynamic programming approach of Askar *et al.* (2007) and ongoing criticism of the standard methods in aggregate planning we developed a new Mixed Integer Linear Programming (MIP) modelling approach for the problem of aggregated production planning in the automotive industry. The modelling especially considers the challenges arising from the possible use of flexibility in workforce and working times. We develop a set of primal heuristics to be able to solve the problem efficiently. A case study and results based on a testset are presented to show the capability of the model and the implementation. As the planning problem addressed by our approach differs in some details, e.g. the consideration of buffers and change costs, from the problem addressed by Askar *et al.* (2007), we do not include a direct comparative study of the two approaches. Based on our experience, we can say that which approach is superior depends on the planning situation and problem instance. The dynamic programming approach has advantages on single-line problems without buffers and change costs whereas the MIP approach is superior on the more complex instances. Furthermore, the MIP approach benefits from easier maintainability and ever improving standard software to solve mixed integer programming problems.

The remainder of this paper is organized as follows. In the following section the problem is described in detail. In section 3 our MIP modelling approach is presented. In sections 4 and 5 implementation details and a case study are reported. Furthermore, our findings regarding the solution time of different problem instances are analysed. The paper ends with a conclusion and an outlook on future research.

2. Aggregate production planning in the automotive industry

2.1. Automotive production

The production process in an automotive plant for car manufacturing contains four main production stages: the stamping of metal or aluminum sheets for the body, the body shop where the body of the car is welded together, the paint shop where the painting and drying of the body is done and the final assembly where the painted body as well as engine, axles, transmission and the interior are assembled together. Each of the stages is decoupled by buffers of limited size. In this paper we consider the production process starting

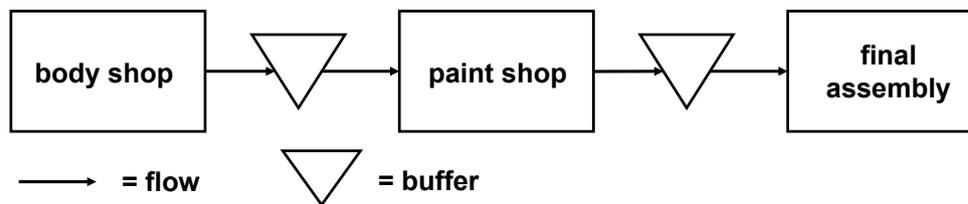


Figure 1. Structure of the manufacturing process

after the stamping plant as it is shown in figure 1. From this point onwards each shop has got one or more flow shop production lines which themselves consist of a number of serially arranged assembly stations, between which cars are conveyed with a fixed belt rate also known as cycle time (see Meyr 2004).

To fulfill market demand the planning starts by adjusting the capacity of the final assembly and hands over the resulting production program consecutively to the preceding shops which are planned under consideration of the buffers. Similar settings for the planning of flow shop lines can be found in the area of truck manufacturing and the production of modules such as engines, axles and transmissions. They have got an identical structure but differ in labor demand, production capacity, etc. (see Askar *et al.* 2007).

2.2. Flexibility

Demand flexibility is the capability of a production system to adapt efficiently to changing demands. The need for demand flexibility in aggregate planning is driven by fluctuations in market demand. They occur for reasons such as the product lifecycle and seasonalities. Decisions regarding the installation of production lines and the product allocation are made in strategic planning. Therefore, for the mid term planning horizon, especially the flexibility of workforce and working time play an important role. The deployment of workforce is organized by means of weekly shift models. A shift model is defined by the number and the length of shifts it contains in a week. A shift belongs to a shift group, i.e., early-, late- or night-shift. In the German automotive industry, the number of different shift groups, the length of the shifts and the number of shifts per week can vary within the year only since the beginning of the 90's. This enhanced flexibility is supported by additional concepts such as working time accounts to lend or borrow working time from employees to compensate demand fluctuations over time. Moreover, temporary workforce plays an important role to adapt to demand. To decouple consecutive production lines from each other and to be able to adapt in each production line in the most efficient way, buffers are used. Technical flexibility may be provided as well in terms of changes in cycle time, the ratio of platforms on the belt that are used or the configuration of different parallel line segments as seen in the paint shop. However, the use of these flexibility measures is sometimes limited by contracts with the labor union or governmental laws. It is important to consider these regulatory requirements to be able to form executable, i.e. feasible, production plans.

2.3. *Market demand and production capacity*

In our problem setting, a dynamic product demand, varying over time, for multiple products is given. In case of the final assembly it is the weekly sales forecast, for paint shop and body shop it is the demand given by the preceding stage on a shiftwise basis. The consideration of a shiftwise production program is important because of the decoupling situation induced by the buffers. Moreover, rescheduling of the production program in the final assembly is possible, which directly effects demand of the preceding stages. Stochastic demand is not considered here due to the fact that the problem complexity would raise dramatically and that information or distributions required are not known or available.

The product demand has to be satisfied by the production capacity. If a decoupling of the production over buffers is possible, the buffers have a minimum and maximum amount of stock. Depending on the specific situation, the minimum stock is either equal to zero or, in case of a safety stock, bigger than zero. The production capacity is determined by the working time and the output per hour. The working time is determined by the selection of a shift model whereas for each week a shift model is determined. Quite often there are regulations which limit the selection and the length of the shift models. Since shifts can still be canceled after a shift model has been chosen, a differentiation is made between theoretic and real working time. The selection of a shift model sets the maximum working time for a shift in this model which is also called the theoretic working time. If the shift is not canceled, i.e., it is actually used for production, it contributes to the real working time. First the shift model is selected for the week, afterwards the decision about the utilization of the individual shifts is made. Furthermore, the output rate of each production stage has to be determined for each week. The output rate is measured in units per hour and each production stage has a set of discrete output possibilities. The combination of shift model and output leads to production capacity. Again a differentiation between theoretic and real production capacity can be made analog to the working time.

The occurring costs are the costs per produced unit and holding costs. Moreover, costs for the selection or change of a shift model and output rate may occur. With regard to the change costs a further differentiation is made between changes among shift models with the same or a different number of shift groups.

2.4. *Workforce demand and workforce capacity*

The workforce demand can be derived from the production capacity planning. It is determined by the combination of shift model and output. There is a basic need for workers directly at the flowshop line which is increased by workers that carry out common tasks such as logistics; together they form the net demand for workers. Based on this value workers for absenteeism and free shift planning are added to calculate the gross demand. The workers for free shift planning are necessary because workers are sometimes eligible for days off when they work longer than it is settled in their working agreement. I. e., if overtime is used in excess in one week, individual days off have to be granted the next week. The number of workers for free shift planning depends on the shift model.

The gross demand for workers has to be satisfied by the group of regular and temporary workers. There is a limit on the proportion of temporary workers in relation to the entire workforce. Moreover the number of workers of each group that are hired or dismissed is limited for each period. The working time account makes it possible to borrow from or

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
lend working time to the employees. The balance of the working time account is calculated by building the difference of working hours according to the working agreements and the actual hours as defined by the real working time. This difference is multiplied by the number of hired workers excluding the workers for absenteeism. These workers do not contribute to the working time account. Finally, an average is built by dividing by the total number of workers.

The occurring costs are wage costs and costs for overtime premiums. These costs are differentiated by workers working directly at the flowshop line and workers doing common tasks. Additionally, there are change costs for hiring and dismissing regular and temporary workers. Note, that there are no costs in the sense of monetary payments associated with the working time account.

2.5. *Current planning process in practise, goals and limitations of a model-based approach*

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Currently, the described planning problem is typically addressed by a continuous monthly planning process with a rolling horizon of fifty-two to seventy-eight weeks. The planning is strongly influenced by the organization structure and therefore conducted sequentially by a number of different departments. At first, a logistics department works out a production program to satisfy the demand by selecting shift models and shifts for production. Afterwards, a workforce planning department calculates the number of needed workers and the development of the working time account. Then a human resources department has to match the demand for workers with the actual number of employees. Finally, all numbers are put together by a productivity department which also conducts a financial evaluation of the costs. Such an organizational structure is found at almost any car maker as these are the standard planning functions necessary for production not only in the light of APP but for many other issues as well. This process can also iterate over the departments evaluating different scenarios in order to find a cost-efficient solution which fulfils the demand. Still the challenge of this process is that each department seeks local optima and that there is no central department with the aim to optimize the overall costs. However, given a model which represents all the necessary requirements and knowledge of the other departments, such a function can be put in place in the productivity department. Thereby, the productivity department is enabled to generate ideas and scenarios for an overall optimum and then discuss them with the other departments responsible for the detailed planning.

The model which will be presented in the remainder of this paper has been developed exactly for this purpose. Among other considerations, it has to take into account union requirements and labour agreements. For instance, lay offs of staff can be forbidden or the dismissal of workers can be associated with compensation payments. Furthermore, the explicit steering of the average working time account with lower and upper bounds should be possible. To increase applicability and acceptance of a model-based approach in practise, the model has to be embedded into a decision support system which eases data input, the generation of alternative solutions for different scenarios, report generation and visualisation of planning solutions.

The limits of such an approach mainly lie within the planning organisation itself. The involved departments should be committed to seek the overall optimum instead of pursuing their own special interests. A model-based approach can help to objectify communication processes and to make decisions more transparent.

3. A mixed-integer linear programming model

The main challenge of modelling the problem is to anticipate the operational shift planning correctly and to gain a solution that can be realized under the consideration of buffers. The main decisions on the maximal production capacity as a combination of output, shift model and the number of workers are made on a weekly basis. The decisions that are made within the week concern the buffers, the working time account and the utilization of the capacity. They are made for each shift to anticipate the operational planning.

We take a more general approach in modelling the problem that is not solely focused on a decomposition in weeks and shifts. There is a differentiation made between macro and micro periods in our model, which however could correspond to weeks or shifts. The main decision inducing costs is made in the macro periods while the planning of micro periods is included to make sure that the solutions are executable on the next planning level.

3.1. Production capacity planning

In our integrated model, a planning horizon of N macro periods each consisting of M micro periods will be considered. With respect to the problem the macro periods are equivalent to weeks whereas the micro periods are equivalent to shifts. There are I combinations of shift models and output referred to as production capacity levels and J products. Each of the I production capacity levels has a subset of micro periods where it provides capacity. The corresponding index sets are defined as follows:

$$\begin{aligned}\bar{N} &= \{1, \dots, N\}, & \bar{M} &= \{1, \dots, M\}, \\ \bar{I} &= \{1, \dots, I\}, & \bar{J} &= \{1, \dots, J\}.\end{aligned}$$

The main task of production capacity planning is to provide enough capacity to satisfy the dynamic demand. The maximum production capacity of a macro period can be calculated in advance depending on the production capacity level i as the product of working time derived from the shift model and the output per hour. In this part of the model we use the following notation:

Parameters

$Y_{n,m,i}^{avl} \in \{0, 1\}$:= Parameter which is equal to 1 if production capacity level i provides capacity in micro period m and macro period n , else 0.

$C_{n,m,i} \in \mathbb{R}^+$:= Production capacity in units at production capacity level i in micro period m and macro period n .

$MCU_n \in [0, 1]$:= Minimum capacity utilization of each micro period that is used in macro period n .

$D_{n,m,j} \in \mathbb{R}^+$:= Demand of product j in units in micro period m and macro period n .

$SS_{n,j} \in \mathbb{R}^+$:= Security stock for product j in macro period n .

$MS_{n,j} \in \mathbb{R}^+$:= Maximum stock for product j in macro period n .

$Q \in \mathbb{N}$:= Smallest number of macro periods for which a production capacity level has to be valid after a change of the production capacity level occurred.

$K^{prod} \in \mathbb{R}^+$:= Production cost per unit.

$K^{stock} \in \mathbb{R}^+$:= Holding costs per unit and micro period.

$K^{micro} \in \mathbb{R}^+ :=$ Fixed costs of a micro period (which typically corresponds to a shift).
 $K_{i,i'}^{change} \in \mathbb{R}^+ :=$ Cost to change from production capacity level i to i' ($i \neq i'$).
 $BIGM \in \mathbb{R}^+ :=$ A constant so large that the constraint cannot be violated if the associated binary variable is 1.

Decision Variables

$\mathbf{y}_{n,i}^{macro} \in \{0, 1\} :=$ Binary variable which is 1 if production stage i is selected in macro period n , else 0.

$\mathbf{y}_{n,m}^{micro} \in \{0, 1\} :=$ Binary variable which is 1 if micro period m is utilized in macro period n , else 0.

$\mathbf{c}_{n,m}^t \in \mathbb{R}^+ :=$ Theoretic production capacity in micro period m and macro period n .

$\mathbf{c}_{n,m}^r \in \mathbb{R}^+ :=$ Real production capacity in micro period m and macro period n .

$\mathbf{x}_{n,m,j} \in \mathbb{R}^+ :=$ Number of produced units of product j in micro period m and macro period n .

$\mathbf{s}_{n,m,j} \in \mathbb{R}^+ :=$ Number of units in stock of product j at the end of micro period m and macro period n

$\mathbf{s}_{0,0,j} =$ starting inventory.

$\mathbf{q}_{n,i,i'} \in \{0, 1\} :=$ Binary variable which is equal to 1 if a change in production capacity from stage i to i' occurred at the beginning of macro period n , else 0 ($i \neq i'$).

The fractional variable $\mathbf{x}_{n,m,j}$ is rounded after the problem is solved to gain integral results. This is sufficient due to the fact that the values are large. The variable $\mathbf{c}_{n,m}^t$ represents a bookkeeping variable and is needed to model an upper bound on the real production capacity in the following set of constraints. The part of production capacity planning is modeled as follows.

$$\begin{aligned}
 \min \mathbf{z}^p = & \sum_{n=1}^N \sum_{m=1}^M \sum_{j=1}^J K^{prod} \cdot \mathbf{x}_{n,m,j} + \sum_{n=1}^N \sum_{m=1}^M \sum_{j=1}^J K^{stock} \cdot \mathbf{s}_{n,m,j} \\
 & + \sum_{n=1}^N \sum_{m=1}^M K^{micro} \cdot \mathbf{y}_{n,m}^{micro} + \sum_{n=1}^N \sum_{i=1}^I \sum_{i' \in \bar{I} \setminus \{i\}} K_{i,i'}^{change} \cdot \mathbf{q}_{n,i,i'}
 \end{aligned} \tag{1}$$

$$\mathbf{c}_{n,m}^t = \sum_{i=1}^I (C_{n,m,i} \cdot \mathbf{y}_{n,i}^{macro}) \quad \forall n \in \bar{N}, \forall m \in \bar{M} \tag{2}$$

$$\mathbf{c}_{n,m}^r \leq \mathbf{c}_{n,m}^t \quad \forall n \in \bar{N}, \forall m \in \bar{M} \tag{3}$$

$$\mathbf{c}_{n,m}^r \leq BIGM \cdot \mathbf{y}_{n,m}^{micro} \quad \forall n \in \bar{N}, \forall m \in \bar{M} \tag{4}$$

$$\sum_{j=1}^J \mathbf{x}_{n,m,j} \leq \mathbf{c}_{n,m}^r \quad \forall n \in \bar{N}, \forall m \in \bar{M} \tag{5}$$

$$\sum_{j=1}^J \mathbf{x}_{n,m,j} \geq MCU_n \cdot \mathbf{c}_{n,m}^r \quad \forall n \in \bar{N}, \forall m \in \bar{M} \tag{6}$$

$$\mathbf{s}_{n,m,j} = \mathbf{s}_{n,m-1,j} + \mathbf{x}_{n,m,j} - D_{n,m,j} \quad \forall n \in \bar{N}, \forall m \in \bar{M}, \forall j \in \bar{J} \tag{7}$$

$$SS_{n,j} \leq s_{n,m,j} \leq MS_{n,m,j} \quad \forall n \in \bar{N}, \forall m \in \bar{M}, \forall j \in \bar{J} \quad (8)$$

$$y_{n-1,i}^{macro} + y_{n,i'}^{macro} \leq 1 + q_{n,i,i'} \quad \forall n \in \bar{N}, \forall i, i' \in \bar{I}, i \neq i' \quad (9)$$

$$\sum_{n=n'}^{n'+Q} \sum_{i=1}^I \sum_{i' \in \bar{I} \setminus \{i\}} q_{n,i,i'} \leq 1 \quad \forall n' \in \bar{N} \text{ and } n' \leq N - Q \quad (10)$$

$$\sum_{i=1}^I y_{n,i}^{macro} = 1 \quad \forall n \in \bar{N} \quad (11)$$

$$y_{n,m}^{micro} + y_{n,i}^{macro} \leq 1 + Y_{n,m,i}^{avl} \quad \forall n \in \bar{N}, \forall m \in \bar{M}, \forall i \in \bar{I} \quad (12)$$

The objective function (1) considers costs for production, holding, the usage of each micro period and changes in capacity. The constraint sets (2) to (4) ensure that the requirements for the real capacity are fulfilled. In this case *BIGM* should be selected as $max(C_{n,m,i})$. The units produced are limited by constraint group (5), which represents the available capacity. The flow shop principle makes no further differentiation among products necessary although it would be possible. The constraint set (6) ensures that if production occurs the minimum utilization rate of a micro period is considered. Constraint set (7) represents the stock equation of the buffer (with $m - 1$ in $s_{n,m-1,j}$ being the preceding micro period) and (8) puts lower and upper bounds on the amount of stock. It is important to consider changes in the production capacity levels. Moreover it is has to be ensured that the number of possible changes can be regulated. The constraint set (9) ensures that changes are indicated by binary variables. By adding upper and lower bounds on the $q_{n,i,i'}$ variable, it is possible to exclude certain changes. In practice, this could be certain combinations of shift models and output or the change from a certain shift model to another. The constraint set (10) prohibits more than one change in the corresponding time-window Q . This accounts for the fact that, in practical applications, changes in the production capacity level may not be wanted in periods not far apart from each other. Two sets of constraints are added to ensure the proper usage and linking of the binary variables. The first group of constraints (11) ensures that exactly one production capacity level is selected whereas (12) forces that only macro periods can be utilized when the decision on the micro period provides capacity for the corresponding period.

The above buffer concept can be used to consider weekly demands, too. In this case, the entire weekly demand needs to be fulfilled in the last micro period of the week and the minimum and maximum stocks need to be adjusted accordingly. It is more efficient to develop a stock equation which is based on the week however. Another challenge arises if the micro periods need to be planned in a certain sequence. Due to the fact that each production capacity level has got its own subset of usable micro periods the constraints become quite complicated, yet it is possible to find a formulation. Both developments are skipped here for reasons of a compact presentation but can be provided upon request.

With regard to the production capacity there is a correlation to lot sizing models. Once production capacity levels are fixed, meaning that values for the $y_{n,i}^{macro}$ variables are set, the resulting structure is similar to a capacitated lot sizing problem. However, the minimum capacity utilization is usually not part of these problems.

3.2. Workforce capacity planning

Workforce capacity planning covers the calculation and satisfaction of workforce demand in terms of the number of regular and temporary workers. Furthermore, changes in workforce capacity and salary premiums have to be considered. The following model fragments extend the above formulation by workforce capacity planning.

Parameters

$A_{n,i}^{line} \in \mathbb{R}^+$:= Demand for workers directly at the line in production capacity level i and macro period n .

$A_{n,i}^{com} \in \mathbb{R}^+$:= Demand for workers doing common tasks in production capacity level i and macro period n .

$A_{n,i}^{net} \in \mathbb{R}^+$:= Net demand for workers in production capacity level i and macro period n .

$A_{n,i}^{net} \in \mathbb{R}^+ := A_{n,i}^{line} + A_{n,i}^{com}$.

$A_n^{Abs} \in [0, 1]$:= Fraction of workers which are absent due to sickness in period n .

$A_{n,i}^{FSP} \in [0, 1]$:= Fraction of workers which are needed for compensating free shifts (days off) of other workers in production capacity level i in period n .

$A_{n,i}^{gross} \in \mathbb{R}^+$:= Gross demand for workers in production capacity level i and macro period n : $A_{n,i}^{gross} := \left[A_{n,i}^{net} \cdot \frac{1}{1 - A_n^{Abs} - A_{n,i}^{FSP}} \right]$.

$A_n^{PTW} \in [0, 1]$:= Maximum ratio of temporary workers from overall workforce in period n .

$AKR_n^{in} \in \mathbb{R}^+$:= Maximum number of regular workers that can be hired for macro period n .

$AKR_n^{out} \in \mathbb{R}^+$:= Maximum number of regular workers that can be dismissed in macro period n .

$AKT_n^{in} \in \mathbb{R}^+$:= Maximum number of temporary workers that can be hired for macro period n .

$AKT_n^{out} \in \mathbb{R}^+$:= Maximum number of temporary workers that can be dismissed in macro period n .

$SP_{n,m,i} \in \mathbb{R}^+$:= Hours relevant for shift premiums in macro period n , micro period m and production capacity level i .

$K_{wage}^{line} \in \mathbb{R}^+$:= Wage of a line worker per macro period.

$K_{wage}^{com} \in \mathbb{R}^+$:= Wage of a common worker per macro period.

$K_{wage}^{tmp} \in \mathbb{R}^+$:= Wage of a temporary worker per macro period.

$K_{m,i}^{prem} \in \mathbb{R}^+$:= Premium per hour in micro period m and capacity level i .

$K_{change}^{reg.in} \in \mathbb{R}^+$:= Costs for hiring a regular worker.

$K_{change}^{reg.out} \in \mathbb{R}^+$:= Costs for dismissing a regular worker.

$K_{change}^{tmp.in} \in \mathbb{R}^+$:= Costs for hiring a temporary worker.

$K_{change}^{tmp.out} \in \mathbb{R}^+$:= Costs for dismissing a temporary worker.

Decision Variables

$\mathbf{a}_n^{gross} \in \mathbb{R}^+$:= Gross demand for workers in macro period n .

$\mathbf{a}_n^{com} \in \mathbb{R}^+$:= Demand for common workers in macro period n .

$\mathbf{a}_n^{reg} \in \mathbb{R}^+$:= Number of regular workers in macro period n

$\mathbf{a}_0^{reg} \in \mathbb{R}^+$:= Start value.

$\mathbf{a}_n^{tmp} \in \mathbb{R}^+$:= Number of temporary workers in macro period n

$\mathbf{a}_0^{tmp} \in \mathbb{R}^+$:= Start value.

$\mathbf{a.in}_n^{reg} \in \mathbb{R}^+ :=$ Number of regular workers hired at the beginning of macro period

$\mathbf{a.out}_n^{reg} \in \mathbb{R}^+ :=$ Number of regular workers dismissed at the beginning of macro period n .

$\mathbf{a.in}_n^{tmp} \in \mathbb{R}^+ :=$ Number of temporary workers hired at the beginning of macro period n .

$\mathbf{a.out}_n^{tmp} \in \mathbb{R}^+ :=$ Number of temporary workers dismissed at the beginning of macro period n .

$\mathbf{sp}_{n,m}^t \in \mathbb{R}^+ :=$ Number of hours theoretically relevant for shift premiums in macro period n and micro period m

$\mathbf{sp}_{n,m}^r \in \mathbb{R}^+ :=$ Number of hours actually relevant for shift premiums in macro period n and micro period m

$$\begin{aligned} \min \mathbf{z}^w = & \sum_{n=1}^N K_{wage}^{line} \cdot \mathbf{a}_n^{reg} + \sum_{n=1}^N \left((K_{wage}^{com} - K_{wage}^{line}) \cdot \mathbf{a}_n^{com} \right) \\ & + \sum_{n=1}^N K_{wage}^{tmp} \cdot \mathbf{a}_n^{tmp} + \sum_{n=1}^N \sum_{m=1}^M K_{m,i}^{prem} \cdot \mathbf{sp}_{n,m}^r \\ & + \sum_{n=1}^N K_{change}^{reg-in} \cdot \mathbf{a.in}_n^{reg} + \sum_{n=1}^N K_{change}^{reg-out} \cdot \mathbf{a.out}_n^{reg} \\ & + \sum_{n=1}^N K_{change}^{tmp-in} \cdot \mathbf{a.in}_n^{tmp} + \sum_{n=1}^N K_{change}^{tmp-out} \cdot \mathbf{a.out}_n^{tmp} \end{aligned} \quad (13)$$

$$\mathbf{a}_n^{com} = \sum_{i=1}^I (A_{n,i}^{com} \cdot \mathbf{y}_{n,i}^{macro}) \quad \forall n \in \bar{N} \quad (14)$$

$$\mathbf{a}_n^{gross} = \sum_{i=1}^I (A_{n,i}^{gross} \cdot \mathbf{y}_{n,i}^{macro}) \quad \forall n \in \bar{N} \quad (15)$$

$$\mathbf{a}_n^{com} \leq \mathbf{a}_n^{reg} \quad \forall n \in \bar{N} \quad (16)$$

$$\mathbf{a}_n^{gross} \leq \mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp} \quad \forall n \in \bar{N} \quad (17)$$

$$\mathbf{a}_n^{tmp} \leq A_n^{PTW} \cdot (\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}) \quad \forall n \in \bar{N} \quad (18)$$

$$\mathbf{a}_n^{reg} = \mathbf{a}_{n-1}^{reg} + \mathbf{a.in}_n^{reg} - \mathbf{a.out}_n^{reg} \quad \forall n \in \bar{N} \quad (19)$$

$$\mathbf{a}_n^{tmp} = \mathbf{a}_{n-1}^{tmp} + \mathbf{a.in}_n^{tmp} - \mathbf{a.out}_n^{tmp} \quad \forall n \in \bar{N} \quad (20)$$

$$\mathbf{sp}_{n,m}^t = \sum_{i=1}^I (SP_{n,m,i} \cdot \mathbf{y}_{n,i}^{macro}) \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (21)$$

$$\mathbf{sp}_{n,m}^r \leq \mathbf{sp}_{n,m}^t \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (22)$$

$$\mathbf{sp}_{n,m}^r \leq BIGM \cdot \mathbf{y}_{n,m}^{micro} \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (23)$$

$$\mathbf{sp}_{n,m}^r - \mathbf{sp}_{n,m}^t + (1 - \mathbf{y}_{n,m}^{micro}) \cdot BIGM \geq 0 \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (24)$$

$$\mathbf{a.in}_n^{reg} \leq AKR_n^{in} \quad \forall n \in \bar{N} \quad (25)$$

$$\mathbf{a.out}_n^{reg} \leq AKR_n^{out} \quad \forall n \in \bar{N} \quad (26)$$

$$\mathbf{a.in}_n^{tmp} \leq AKT_n^{in} \quad \forall n \in \bar{N} \quad (27)$$

$$\mathbf{a.out}_n^{tmp} \leq AKT_n^{out} \quad \forall n \in \bar{N} \quad (28)$$

In workforce capacity planning, costs are incurred for regular wages, salary premiums and for changes in workforce. In the objective function (13), workers for common tasks are part of the regular workers and receive a regular wage plus an extra of $(K_{wage}^{com} - K_{wage}^{line})$ for special qualifications. Constraint sets (14) and (15) determine the demand for common workers and the gross demand for workers in dependency of the selected production capacity level, respectively. The inequalities (16) ensure that the demand for common workers is satisfied by regular workers, while constraint set (17) ensures overall demand fulfillment of regular and temporary workers. The set of constraints (18) ensures that the number of temporary workers make up only a certain proportion of the overall workforce. Hiring and dismissal of workers is modeled by the constraint sets (19) to (28). Constraint sets (19) and (20) represent the balance constraints for regular and temporary workers, respectively. Salary premiums play an important role for the costs structure. They occur when micro periods are utilized for production. Constraint sets (21) and (22) ensure that the theoretic number of hours relevant for salary premiums are calculated and placed as an upper bound on $\mathbf{sp}_{n,m}^r$ afterwards. Inequalities (23) and (24) make sure that $\mathbf{sp}_{n,m}^r$ equals $\mathbf{sp}_{n,m}^t$ if the micro period m is utilized. In this case *BIGM* should be selected as $\max(SP_{n,m,i})$. Upper bounds on the changes in workforce are considered in constraint sets (25) to (28).

Finally, the two objective functions (1) and (13) are joined into one overall objective function which is displayed in equation (29).

$$\mathbf{z}^{sum} = \mathbf{z}^p + \mathbf{z}^w \quad (29)$$

3.3. Working time account

The idea of a working time account is to borrow from or lend working time to employees. Thereby, it is possible to level out seasonal variations in working time. In the following, a formulation for modelling an average working time account as described in section 2.4 is presented.

Parameters

$Z_{n,m,i}^{avg} \in \mathbb{R}^+$:= Average working time in micro period m , macro period n and production capacity level i .

$V_n \in \mathbb{R}^+$:= Working time according to tariff agreements in macro period n .

$WTA_n^{LB} \in \mathbb{R}^-$:= Lower bound of the working time account in macro period n .

$WTA_n^{UB} \in \mathbb{R}^+$:= Upper bound of the working time account in macro period n .

Decision Variables

$\mathbf{z}_{n,m}^t \in \mathbb{R}^+$:= Theoretical average working time in macro period n and micro period m .

$\mathbf{z}_{n,m}^r \in \mathbb{R}^+$:= Real average working time in macro period n and micro period m .

$\mathbf{b}_n \in \mathbb{R}$:= Level of the working time account at the end of macro period n .

\mathbf{b}_0 = Level at the beginning.

The negativity of WTA_n^{LB} reflects the fact that the lower bound of an average working

time account may take negative values if working time is borrowed from the employees. This occurs when more working time is used than settled in the tariff agreement. The parameter $Z_{n,m,i}^{avg}$ for the average working time is calculated by building the ratio of workers in the corresponding micro period to the overall net demand for workers and multiplying this ratio with the working time in the micro period. For instance suppose that, for a specific production capacity level, one macro period lasted 8 hours and a total demand of 200 workers had to be satisfied for the macro period. If 100 workers were needed in each micro period that is used, while the remaining 100 workers had days off / free shifts, then $Z_{n,m,i}^{avg}$ would be 4 hours. The decisions are made about average working time and the level of the working time account.

$$\mathbf{z}_{n,m}^t = \sum_{i=1}^I \left(Z_{n,m,i}^{avg} \cdot (1 - A_n^{Abs}) \cdot \mathbf{y}_{n,i}^{macro} \right) \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (30)$$

$$\mathbf{z}_{n,m}^r \leq \mathbf{z}_{n,m}^t \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (31)$$

$$\mathbf{z}_{n,m}^r \leq BIGM \cdot \mathbf{y}_{n,m}^{micro} \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (32)$$

$$\mathbf{z}_{n,m}^r - \mathbf{z}_{n,m}^t + (1 - \mathbf{y}_{n,m}^{micro}) \cdot BIGM \geq 0 \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (33)$$

$$\mathbf{b}_n = \mathbf{b}_{n-1} +$$

$$\sum_{m=1}^M \mathbf{z}_{n,m}^r - (1 - A_n^{Abs}) \cdot V_n \cdot \left[1 + \left(1 - \frac{\mathbf{a}_n^{gross}}{\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}} \right) \right] \quad \forall n \in \bar{N} \quad (34)$$

$$WTA_n^{LB} \leq \mathbf{b}_n \leq WTA_n^{UB} \quad \forall n \in \bar{N} \quad (35)$$

The constraint set (30) calculates the theoretic working time. It is important to note that the fraction of workers who are absent do not contribute to the theoretic working time. The constraint sets (31) to (33) force $\mathbf{z}_{n,m}^r$ to equal $\mathbf{z}_{n,m}^t$ if micro period m in macro period n is used for production and to be zero, otherwise. In this case $BIGM$ should be selected as $\max(Z_{n,m,i}^{avg})$. Constraint set (35) places an upper and lower bound on the working time account. The working time account is updated by constraint set (34). To achieve this the difference of the actually occurred working time and the working time according to tariff is computed and added to the working time account balance of the preceding macro period. The factor $(1 - A_n^{Abs})$ considers the fact that absent workers do not contribute their tariff working time to the overall calculation. The nonlinear factor of

$$0 \leq 1 - \frac{\mathbf{a}_n^{gross}}{\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}} \leq 1 \quad (36)$$

is the ratio of how many workers are hired more than are actually required in terms of gross demand. If it is strictly greater than zero and the difference computed in equation set (34) is negative, this is equivalent to decreasing the working time account by planning a free shift. In the appendix of this paper, we present a linear approximation of ratio (36).

As mentioned before the working time account is, as such, not associated with costs in the sense of monetary payments. In practice, the companies currently use the working time account as a flexibility measure to cope with short time fluctuations and effects of changing demand over the product lifecycle. It only comes into focus when it runs against one of its borders, especially the upper one. In such a case, the decision maker

actively steers the working time account by introducing tighter limits or further constraints. Therefore, from the decision maker's point of view, it can be treated as a pure feasibility requirement. The planning departments typically want the working time account to assume a special value or remain within a certain range independently of the direct monetary effect.

However, due to the finite planning horizon, the usage of the working time account can induce cost effects in subsequent planning periods. From a modelling perspective, there are in principle several ways to deal with this phenomenon. One is to choose a sufficiently long planning horizon such that the possibly distorting effects at the end of the models planning horizon do only have a marginal influence on the planning solution at the beginning. In our case, the decision maker is actually interested in a production and workforce plan for the next 3 - 6 months. The model is solved monthly with a rolling planning horizon of typically 18 to 24 months. Another option would be to consider the working time account in the goal function. This can be done by keeping a separate balance of the hours worked by all workers of the net demand $A_{n,i}^{net}$ and the workers actually employed $\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}$ and their tariff working time V_n . This balance can then be multiplied by a cost factor in the objective function. To do so we introduce the following definitions and constraints.

Parameters

$Z_{n,m,i}^{sum} \in \mathbb{R}^+ :=$ Sum of the working time in micro period m , macro period n and production capacity level i .

$K_{wta}^{hour} \in \mathbb{R}^+ :=$ Cost of one hour in the summed working time account.

Decision Variables

$\mathbf{z}_{n,m}^{t_sum} \in \mathbb{R}^+ :=$ Theoretical average working time in macro period n and micro period m .

$\mathbf{z}_{n,m}^{r_sum} \in \mathbb{R}^+ :=$ Real average working time in macro period n and micro period m .

$\mathbf{b}_n^{sum} \in \mathbb{R} :=$ Level of the working time account in terms of the sum of total working hours at the end of macro period n .

$\mathbf{b}_0^{sum} =$ Level at the beginning.

$$\mathbf{z}_{n,m}^{t_sum} = \sum_{i=1}^I (Z_{n,m,i}^{sum} \cdot \mathbf{y}_{n,i}^{macro}) \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (37)$$

$$\mathbf{z}_{n,m}^{r_sum} \leq \mathbf{z}_{n,m}^{t_sum} \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (38)$$

$$\mathbf{z}_{n,m}^{r_sum} \leq BIGM \cdot \mathbf{y}_{n,m}^{micro} \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (39)$$

$$\mathbf{z}_{n,m}^{r_sum} - \mathbf{z}_{n,m}^{t_sum} + (1 - \mathbf{y}_{n,m}^{micro}) \cdot BIGM \geq 0 \quad \forall n \in \bar{N}, \forall m \in \bar{M} \quad (40)$$

$$\mathbf{b}_n^{sum} = \mathbf{b}_{n-1}^{sum} + \sum_{m=1}^M \mathbf{z}_{n,m}^{r_sum} - (1 - A_n^{Abs}) \cdot V_n \cdot (\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}) \quad \forall n \in \bar{N} \quad (41)$$

The constraint sets (37) to (41) work analogous to the sets described in (30) to (34) but now in the context of the summation of working times and the times settled in the tariff agreement. If desired \mathbf{b}_n^{sum} could be bounded similarly to constraint set (35). The goal function would be adapted as shown in equation (42).

$$\mathbf{z}^{sum} = \mathbf{z}^p + \mathbf{z}^w + K_{wta}^{hour} \cdot \mathbf{b}_n^{sum} \quad (42)$$

1
2
3
4
5 A difficulty of the latter approach is, that it is not clear how to determine reasonable
6 values for the cost coefficients K_{wta}^{hour} since they depend on the further development of
7 demand and production after the end of the planning horizon. Intuitively, the costs can
8 be used to prevent the model to consume a “free lunch” by emptying the working time
9 account at the end of the planning horizon. In the framework of a decision support ap-
10 proach the planner can play with these coefficients to generate more consistent solutions.
11 In practise, however, we have mostly seen decision makers neglecting these costs and
12 actively steering just the level of the working time account at the end of the planning
13 horizon by using lower and upper bounds.
14

15 16 17 18 4. Solution framework and integration into a decision support system 19

20 Our research has shown that some practical instances of the above model could not be
21 solved in reasonable time and sufficient solution quality by a standard state-of-the-art
22 solver (cp. results in section 5.2). The performance could be increased by developing
23 specialized model preprocessing, model specific valid inequalities and primal heuristics.
24 In the following, we want to briefly scetch the main ideas. Detailed descriptions can be
25 found in Sillekens (2009).
26

27 The specialized preprocessing uses demand information that is available in advance
28 to fix subsets of the binary variables. If a production capacity stage cannot satisfy the
29 demand minus the maximum stock for a macro period, considering that all micro periods
30 are used, the corresponding capacity stage cannot be used and therefore the binary
31 variable can be fixed to zero. Micro periods can be fixed to one in some cases, e.g, if the
32 minimum stock requirements are violated because even maximum stock cannot satisfy
33 the entire demand that occurs for the micro period without falling below safety stock.
34 Furthermore, valid inequalities can be added. Looking at the demand and the capacities
35 that a production capacity stage provides, a lower bound on the number of micro periods
36 can be determined. This is true for an overall minimum number of shifts over all capacity
37 stages as well as for each stage individually.
38

39 To be able to deploy primal heuristics, a small framework was developed. It consists
40 of a part which generates solution spaces and one which does a search with different
41 approaches on the solution space. The original solution space is generated by the model
42 and the improvements described so far. Two more solution spaces can be generated. The
43 first one we call the shift model-exclusion solution space. It contains only those capacity
44 levels which by means of shift groups are within some range of the demand. In other
45 words, if demand can be satisfied by at least one shift model with two shift groups, all
46 shift models with three shift groups are excluded. The second solution space is called de-
47 mand window solution space. It accepts parameters *lowernumber* and *uppernumber*. The
48 demand is taken and the *lowernumber* capacity levels that are closest to demand from
49 below as well as the *uppernumber* capacity levels that are closest to demand from above
50 are allowed as the only capacity levels in the solution space. It is assumed that the capac-
51 ity of the level is used fully by utilizing all available micro periods. The determination of
52 this demand window is done for each macro period. Furthermore, three heuristics were
53 developed, which operate on the solution space. The first one is an enhanced LP-and-
54 Fix heuristic. The basic LP-and-Fix procedure is described e. g. by Pochet and Wolsey
55 (2006). It is based on the idea to fix binary variables that are in some ϵ -environment to
56 their integral value in the LP-Relaxation. The original idea is to fix only values that are
57 integral. The second one is the Relax-and-Fix heuristic as discussed by Stadtler (2003)
58
59
60

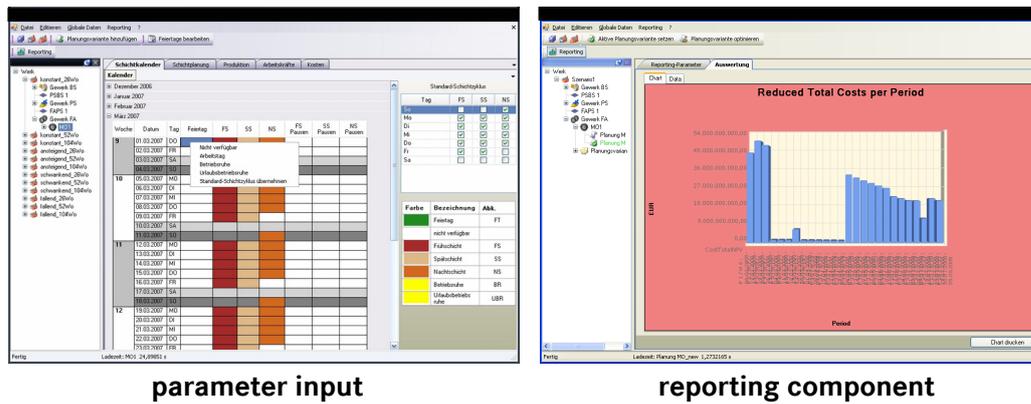


Figure 2. Prototype of the decision support system (illustrative data).

and Pochet and Wolsey (2006), with a timewise decomposition along the macro periods. The last one has, to the best of our knowledge, not been presented in the literature and is problem specific. We call it linking heuristic. Here, the $y_{n,i}^{macro}$ variables are forced to take the same values in subsequent macro periods if the demand of the periods lies only a specified percentage apart. The heuristic is enhanced by a parameter which specifies the maximum number of linked macro periods. Each of the solution spaces can be used together with each heuristic. This guarantees a flexible framework to solve different kinds of problems.

The framework is deployed in a prototype of a decision support system for aggregate production planning at our industrial partner. Our notion of a decision support system (DSS) corresponds to the definition given by Turban and Aronson (1998). It typically contains a data component, a dialogue component and a model or method component, the latter of which corresponds to the solution framework and model just presented above. The data component contains different scenarios that are to be optimized or have to be compared against each other. The dialogue component is crucial for the acceptance of the system. It is subdivided into a data input part, where parameters for the model are maintained, a part for the selection and configuration of the solution method and a reporting and analysis part. A solution generated by the system can be adopted by hand and can be further modified if necessary. Each solution, automatically generated and modified or developed by manual input, can be evaluated and displayed in the reporting component. The reporting component displays all important key indicators for a solution, e.g., the development of the number of regular workers over time. Figure 2 gives an impression of the developed prototype of the decision support system. As the mathematical model as well as the solution algorithm is totally embedded into the system and, hence, hidden from the planner, no in depth knowledge of these components is necessary to use the system in a beneficial way. However, especially in the first weeks of usage a direct and frequent contact to one of the developers turned out to be necessary to introduce and accustom the planner to the notion of model-based planning and adopt the model and the system further to practical requirements.

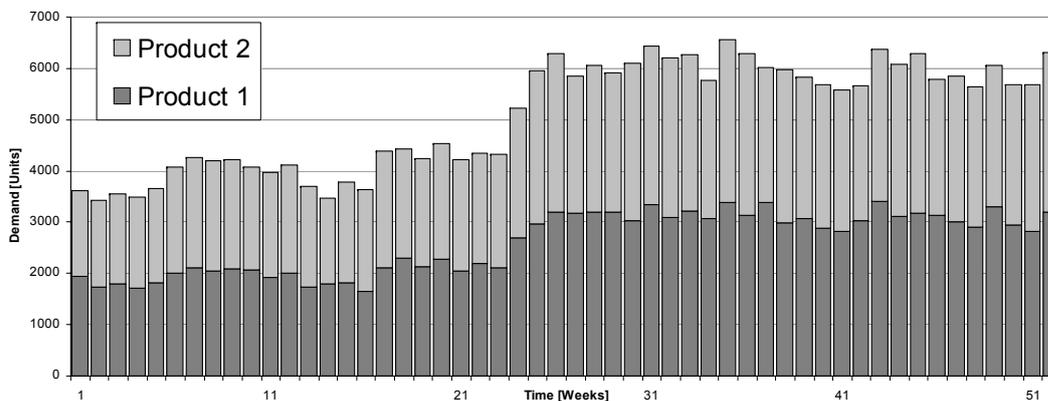


Figure 3. Demand development for two products over time (illustrative data).

5. Numerical results and case study

5.1. Case study

In this section we present an illustrative case study which is supposed to show the applicability of our model and highlight the main decision aspects of an integrated production and workforce planning. In particular, the cost saving potential of workforce flexibility is analysed by investigating how adaptation to demand takes place with different flexibility portfolios. We have to remark, that, due to a non-disclosure agreement we cannot use and publish real world data. However, the parameter values used in this study are realistic in the sense that they represent a typical planning situation in the automotive industry. For the case study a planning horizon of 52 weeks is considered. In each week seven days with up to three shifts are available. In terms of the model this corresponds to 52 macro periods and 21 micro periods within each macro period. An illustrative demand is given for two products and each macro period, it is depicted in Figure 3.

An overall absenteeism of 8% has to be taken into consideration. Furthermore, an additional 50% of workers for common tasks have to be added to the workers employed directly at the line. The working time is seven hours, the value of the working time account at the beginning of the first period is zero. The production capacity levels i in the model result as a product of output rate and shift models. Two flexibility portfolios are available called HighFlex and LowFlex. They consist of shift models and usable output rates and are listed in Table 1. The shift models are named n -shift models in the table with n being the number of shift groups, each separate shift model is written as $n \times PT$ with PT being the available time for production. The actual working time, which is paid for in one shift, is 25 minutes more due to breaks. Production can take place from Monday to Friday, while shift models with one shift group work only early shifts, with two shift groups late shifts are used and three shift groups additionally utilizing a night shift.

There are further parameters for the optimisation of the usage of the two flexibility portfolios. Output rates can be altered earliest twelve weeks after the last change, shift models four weeks after the last change. The overall number of regular workers has to move between 1000 and 3500 with a maximum of 6% temporary workers. We assume, that hiring of workers is always possible as well as the dismissal of temporary workers. At most ten regular workers are allowed to be dismissed each week. The working time account has to be within the bounds of -75 and 75 hours during the entire planning

Parameter	HighFlex	LowFlex
one-shift models	1x6.5h; 1x7.0h; 1x7.5h 1x8.0h; 1x8.5h	1x7.0h; 1x7.5h; 1x8.0h
two-shift models	2x6.5h; 2x7.0h; 2x7.5h 2x8.0h; 2x8.5h	2x7.0h; 2x7.5h; 2x8.0h
three-shift models	3x6.5h; 3x7.0h; 3x7.5h	3x7.0h; 3x7.5h
output	48 jobs/h ($A_{n,i}^{line}=340$) 60 jobs/h ($A^{line}=425$)	60 jobs/h ($A^{line}=425$)

Table 1. Illustrative flexibility portfolios.

Parameter	Value [euro]	Comment
K^{prod}	100	Production of one unit.
K^{micro}	5 000	Usage of one shift.
$K_{i,i'}^{change}$		Change costs depend on type of capacity levels i and i' . A change of the output rate costs 100 000 euro, a change of the shift model 50 000 euro and a change of the shift groups 250 000 euro.
K_{wage}^{line}	1 250	Wage of worker in line per week.
K_{wage}^{com}	1 500	Wage of worker common tasks per week.
K_{wage}^{tmp}	1 000	Wage of worker temporary per week.
$K_{m,i}^{prem}$		The value depends on which shifts are used. Average costs per hour and worker are 35. Overtime premiums are 5% in early shifts, 10% in late shifts and 25% in nightshifts.
$K_{change}^{reg.in}$	2 500	Costs for hiring regular worker.
$K_{change}^{reg.out}$	10 000	Costs for dismissing regular worker.
$K_{change}^{tmp.in}$	500	Costs for hiring temporary worker.
$K_{change}^{tmp.out}$	100	Costs for dismissing temporary worker.

Table 2. Cost parameters for case study (illustrative data).

horizon. Illustrative cost values are shown in Table 2. If costs are not listed it is assumed that they will take a zero value.

In the results of the optimisation, the LowFlex case incurs a total cost of 164 055 010 euro and a working time account of 14.6h in the last period. Regarding the use of the flexibility instruments there is only one output rate at 60 jobs/h which is used in every week. The shift model for the first 24 periods is a 2x8.0h shift model and afterwards a 3x7.5h shift model. The number of workers ranges around 1 400 for the first 24 periods and around 2 100 afterwards. Small fluctuations in workers demand are leveled out by temporary workers. There are no workers employed above the actual gross demand. The change costs make up 2 161 200 euro of the total costs.

The HighFlex case comes out with a total cost of 161 910 463 euro and a working time account of 58.6h in the last period. Regarding the use of the flexibility instruments there are two output rates which are both selected. The output rate of 48 jobs/h is chosen for the first 16 weeks. The 60 jobs/h output rate is selected for the remaining time. The shift model for the first 24 periods is a 2x8.5h shift model and afterwards a 3x7.5h shift model. The number of workers ranges around 1 250 for the first 16 periods, then around 1 400 for eight periods and around 2 100 afterwards. Small fluctuations in worker demand are leveled out by temporary workers. In the first sixteen periods there are about 150 workers employed over the actual gross demand which leads to a smaller increase in the working time account in the beginning. The change costs make up 2 315 700 euro of the

	constant demand scenario around 2-shift model
	rising demand scenario around 2-shift model → 3-shift model
	falling demand scenario around 3-shift model → 2-shift model
	fluctuating demand scenario around 2-shift model → 3-shift model → 2-shift model

Table 3. Illustrative demand scenarios.

total costs.

The portfolios are used similarly with regard to the length of which of the shift models are chosen. This is most likely induced by the fact that the costs for changing a shift model are rather high with 50 000 euro. However, the HighFlex portfolio uses a 2x8.5h shift model in combination with a lower output rate of 48 jobs/h for the first 16 weeks, only afterwards the higher output rate is selected. In both portfolios it is optimal to use a combination of a 3x7.5h shift model with 60 jobs/h for the second half of the planning horizon. The costs for the HighFlex portfolio are 1.8% less than the ones for the LowFlex portfolio. However, the HighFlex portfolio makes use of the working time account finishing 44 hours higher in comparison to the LowFlex case. The HighFlex scenario is, viewed from a cost point, more advantageous and makes more use of the working time account.

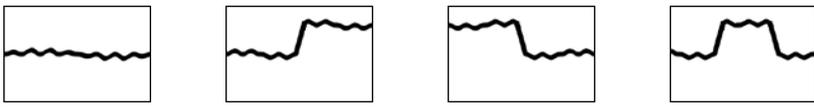
5.2. Computational results

The benchmarking for this section was conducted on a PC with a Windows XP SP2 operating system, 3.2Ghz Dual Core processor and 1GB Ram. To implement the model and to solve problem instances, ILOG CPLEX 11.0 together with ILOG Concert technology has been used.

In the following, the HighFlex portfolio is used on four typical demand scenarios as well as three different planning horizons. The planning horizons are 26, 52 and 104 weeks. The buffer remains at 500 units. The four different demand scenarios are shown in Table 3. The demand is generated randomly within a range of $\pm 10\%$ based on the output quantities of shift models with 7.5h working time. The number of shift groups is taken from Table 3.

The runtime is restricted to two hours. Either both hours are spend on solving the original model, or one hour might be spend on a primal heuristic and the other hour is spend on the original model together with any solution the heuristic might have provided. In Table 4 results are reported for each planning horizon and demand scenario for both solution processes. Results of the solution process without heuristics are reported in the rows marked with the keyword *none*. For the solution process using a primal heuristic only the best result among all the different heuristics is reported in the rows marked with the keyword *best*.

The various solution spaces are abbreviated as follows: n=normal, s=shiftmodel exclu-



WK	Meth.	constant	rising	falling	fluctuation
26	none	12.67 %	5.46 %	2.92 %	0.20 %
	best	11.89 % (d/rf)	5.54 % (d/rf)	1.86 % (n/lp)	0.09 % (d/lp)
52	none	no sol.	4.00 %	1.85 %	3.95 %
	best	9.41 % (d/lh)	3.77 % (s/lh)	1.51 % (s/lp)	3.20 % (d/lh)
104	none	no sol.	no sol.	0.58 %	4.10 %
	best	9.48 % (d/rf)	4.41 % (d/rf)	0.50 % (d/lh)	3.87 % (s/lp)

Table 4. Duality gap for different demand scenarios

sion, d=demand window. Abbreviations for heuristics are: lp=LP-and-Fix, rf=Relax-and-Fix, lh=Linking heuristic. In the case of the heuristic results the most successful combination is put in parenthesis under the *best* solution.

The results in Table 4 show that it is important to have efficient heuristics. Without heuristics, some of the instances could not be solved within the time limit. In all but one case the use of heuristics leads to better results. Regarding the occurring duality gaps, eight out of twelve problems are solved with a gap of less than 5%. Regarding the best solution spaces for the heuristics, the demand window solution space is the best solution space in eight of twelve cases followed by the shift model exclusion solution space in three cases. Among the heuristics, each of the three heuristics is the best in four cases. However, further results, which are not included in the table, show that the LP-and-Fix heuristic often finds no solution at all because it produces an infeasible solution by fixing binary variables.

Further research has to be conducted to get better duality gaps or to solve the problems to optimality. For the solution spaces the demand window and the shift model-exclusion solution space are promising. For the heuristics the Relax-and-Fix heuristic and the linking heuristic are good choices for obtaining upper bounds on the problem.

6. Conclusion and outlook

In this paper we presented a new modelling approach for the problem of aggregate production planning in the automotive industry. In particular, requirements from workforce and working time flexibility were considered. Improvements of the model and the implementation of different primal heuristics were discussed. An illustrative case study was presented which highlights main aspects of integrated capacity and workforce planning. Furthermore, computational results on different problem instances were presented, which show that problem instances of practical size can be solved with sufficient solution quality by a combination of standard solvers and problem specific heuristics and preprocessing. Future work should comprise an integration of different production lines within one plant to increase the potential of flexibility.

Acknowledgements

We wish to thank Dr. Gerhard Jünemann and Dr. Thomas Sommer-Dittrich from Daimler Group Research for their help and support. Furthermore, we thank the four anonymous referees who, by their numerous remarks and suggestions, contributed greatly to improving the manuscript.

References

- Askar, G., *et al.*, 2007. Flexibility Planning in Automotive Plants. *In*: H.O. Günther, D.C. Mattfeld and L. Suhl, eds. *Management logistischer Netzwerke*. Physica, 235–255.
- Askar, G. and Zimmermann, J., 2006. Optimal Usage of Flexibility Instruments in Automotive Plants. *In*: K.H. Waldmann and U.M. Stocker, eds. *Operations Research Proceedings 2006*, Operations Research Proceedings, 479–484.
- Bauer, F., *et al.*, 2007. Cross-country Comparison of Operating Hours, Capacity Utilisation, Working Times and Employment. *In*: L. Delsen, D. Bosworth, H. Groß and R. Muñoz de Bustillo y Llorente, eds. *Operating Hours and Working Times*. Physica, 41–72.
- Becker, H., 2006. *High Noon in the Automotive Industry*. Springer, Berlin.
- Buxey, G., 2003. Strategy not tactics drive aggregate planning. *International Journal of Production Economics*, 85, 331–346.
- Buxey, G., 2005. Aggregate planning for seasonal demand: reconciling theory with practice. *International Journal of Operations & Production Management*, 25 (11), 1083–1100.
- Cha, C.N. and Hwang, H., 1996. Experimental Comparison of the Switching Heuristics for Aggregate Production Planning Problem. *Computers /& Industrial Engineering*, 31 (3/4), 625–630.
- Corominas, A., Lusa, A., and Pastor, R., 2007. Using a MILP model to establish a framework for an annualised hours agreement. *European Journal of Operations Research*, 177, 1495–1506.
- Gilgeous, V., 1989. Modelling realism in aggregate planning: a goal-search approach. *International Journal of Production Research*, 27 (7), 1179–1193.
- Hanssmann and Hess, 1960. A Linear-Programming Approach To Production and Employment Scheduling. *Management Technology*, 1 (1), 110–115.
- Holt, Charles, C., Modiglian, F., and Muth, J.F., 1956. Derivation of a Linear Decision Rule for Production and Employment. *Management Science*, 2 (2), 159–177.
- Holt, Charles, C., Modigliani, F., and Simon, H.A., 1955. A Linear Decision Rule for Production and Employment Scheduling. *Management Science*, 2 (1), 1–30.
- Leung, S.C.H. and Wu, Y., 2004. A robust optimization model for stochastic aggregate production planning. *Production Planning & Control*, 15 (5), 502–514.
- Meyr, H., 2004. Supply chain planning in the German automotive Industry. *OR Spektrum*, 26, 447–470.
- Nam, S. and Logendran, R., 1992. Aggregate production planning - A survey of models and methodologies. *European Journal of Operational Research*, 61, 255–272.
- Nam, S. and Logendran, R., 1995. Modified Production Switching Heuristics for Aggregate Production Planning. *Computers /& Operations Research*, 22 (5), 531–541.
- Ning, Y., Tang, W., and Zhao, R., 2006. Multiproduct Aggregate Production Planning in Fuzzy Random Environments. *World Journal of Modelling and Simulation*, 2 (5),

- 312–321.
- Oloff, M.D., Lewis, H.S., and Markland, R.E., 1989. Aggregate Planning in Crew-Loaded Production Environments. *Computers & Operations Research*, 16 (1), 13–25.
- Orr, D., 1962. A Random Walk Production-Inventory Policy: Rationale and Implementation. *Management Science*, 9, 108–122.
- Pochet, Y. and Wolsey, L.A., 2006. *Production Planning by Mixed Integer Programming*. Springer Series in Operation Research and Financial Engineering Springer, Berlin.
- Sillekens, T., 2009. Aggregierte Produktionsplanung in der Automobilindustrie unter besonderer Berücksichtigung von Personalflexibilität. Thesis (PhD). University of Paderborn.
- Silva, C., *et al.*, 2006. An interactive Decision Support System for an Aggregate Production Planning Model Based on Multiple Criteria Mixed Integer Linear Programming. *Omega - The International Journal of Management Science*, 34 (2), 167–177.
- Stadtler, H., 2003. Multilevel lot-sizing with setup times and multiple constrained resources: Internally rolling schedules with lot-sizing windows. *Operations Research*, 53, 487–502.
- Taubert, W.H., 1968. A Search Decision Rule for the Aggregate Scheduling Problem. *Management Science*, 14 (6), 343–359.
- Turban, E. and Aronson, J.E., 1998. *Decision Support Systems and Intelligent Systems*. 5th edition. Prentice-Hall, Inc.
- Wang, R.C. and Lang, T.F., 2005. Applying possibilistic linear programming to aggregate production planning. *International Journal of Production Economics*, 98, 328–341.

Appendix A. Linear approximation of the working time account

This is a brief description of a linear approximation of ratio (36). For a more comprehensive presentation including detailed derivations and an error analysis we refer to Sillekens (2009).

Let $\mathbf{akv}_{n,i}$ denote the ratio of workers that are employed above gross demand in macro period n and production capacity level i . If capacity level i is chosen in macro period n , $\mathbf{akv}_{n,i}$ can be computed as follows:

$$\mathbf{akv}_{n,i} = 1 - \frac{\mathbf{a}_n^{\text{gross}}}{\mathbf{a}_n^{\text{reg}} + \mathbf{a}_n^{\text{tmp}}} \quad \text{if } \mathbf{y}_{n,i}^{\text{macro}} = 1 \quad (\text{A1})$$

If capacity level i is not chosen in macro period n , $\mathbf{akv}_{n,i}$ is zero. To gain enough precision a series of tangents is required to approximate the relationship for different intervals of the x-axis, i.e., the number of employees ($\mathbf{a}_n^{\text{reg}} + \mathbf{a}_n^{\text{tmp}}$) (see Figure A1). Our research has shown that the approximation by only one tangent is not sufficient to represent changes in the shift model correctly, when there is a big shortage or surplus of workers. A number of four tangents proved to be suitable in our case. If one opts to use more than four tangents the osculation points most certainly need an adjustment. With these preliminaries, the linear approximation can be implemented into the model as follows.

Parameters

$K \in \mathbb{N} :=$ number of approximation intervals (= number of tangents), with corresponding index set $\bar{K} = \{1, \dots, K\}$

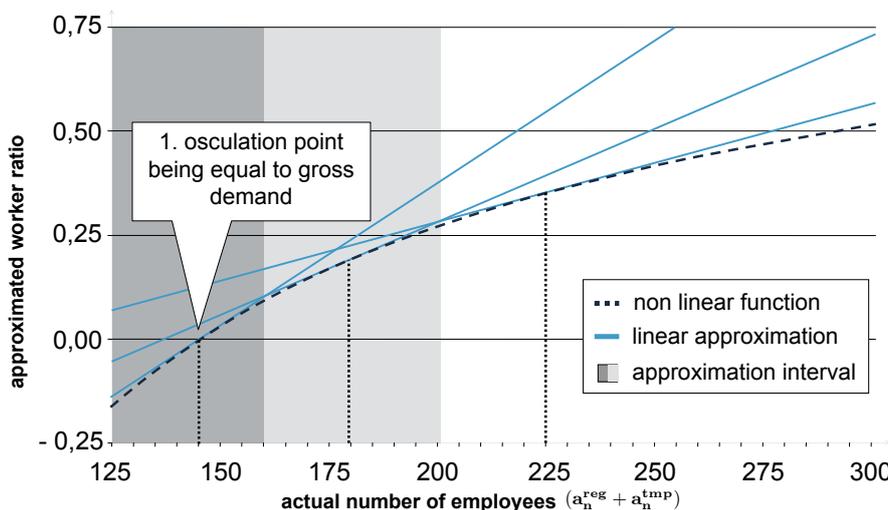


Figure A1. Linear approximation of worker ratio (illustrative).

$LA_{n,i,k}^{TanPt} \in \mathbb{R}^+ :=$ Osculation point of the tangent for approximation interval k for the linear approximation of worker ratio in macro period n for production capacity level i . The function to calculate $LA_{n,i,k}^{TanPt}$ is based on experience values. It is possible to use different points: $LA_{n,i,1}^{TanPt} := A_{n,i}^{gross}$, $LA_{n,i,k}^{TanPt} := LA_{n,i,k-1}^{TanPt} + A_{n,i}^{net} \cdot \frac{1}{1-A_{n,i}^{Abs}} \cdot k$

$LA_{n,i,k}^{grad} \in \mathbb{R}^+ :=$ Gradient of the tangent for approximation interval k for the linear approximation of worker ratio in macro period n for production capacity level i : $LA_{n,i,k}^{grad} := \frac{A_{n,i}^{net} \cdot \frac{1}{1-A_{n,i}^{Abs} - A_{n,i}^{FSP}}}{(LA_{n,i,k}^{TanPt})^2}$

$LA_{n,i,k}^Y \in \mathbb{R} :=$ y-axis intercept of the tangent in approximation interval k for the linear approximation of worker ratio in macro period n for production capacity level i : $LA_{n,i,k}^Y \in \mathbb{R} := 1 - \frac{2 \cdot LA_{n,i,1}^{TanPt}}{LA_{n,i,k}^{TanPt}}$

$LA_{n,i,k}^{AB} \in \mathbb{R}^+ :=$ Left border of the approximation interval k for the linear approximation of worker ratio in macro period n for production capacity level i : $LA_{n,i,1}^{AB} := 0$, $LA_{n,i,k}^{AB} := \frac{LA_{n,i,k}^Y - LA_{n,i,k-1}^Y}{LA_{n,i,k-1}^{grad} - LA_{n,i,k}^{grad}}$

Decision Variables

$akv_{n,i} \in [0, 1] :=$ Ratio of workers that are employed above gross demand in macro period n and production capacity level i .

$y_{n,i,k}^{apx} \in \{0, 1\} :=$ Binary variable which is 1 when the approximation interval k has to be used in macro period n and production capacity level i , else 0.

$$(a_n^{reg} + a_n^{reg}) \leq LA_{m,i,k+1}^{AB} + BIGM \cdot (1 - y_{n,i,k}^{apx}) \quad \forall n \in \bar{N}, \forall i \in \bar{I}, \forall k \in \bar{K} \setminus \{K\} \quad (A2)$$

$$(a_n^{reg} + a_n^{tmp}) \geq LA_{m,i,k}^{AB} - BIGM \cdot (1 - y_{n,i,k}^{apx}) \quad \forall n \in \bar{N}, \forall i \in \bar{I}, \forall k \in \bar{K} \quad (A3)$$

$$\sum_{k=1}^K \mathbf{y}_{n,i,k}^{apx} = 1 \quad \forall n \in \bar{N}, \forall i \in \bar{I} \quad (\text{A4})$$

$$\mathbf{akv}_{n,i} \leq \left(\begin{array}{l} LA_{n,i,k}^Y + LA_{n,i,k}^{grad} \cdot (\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}) + \\ BIGM \cdot (1 - \mathbf{y}_{n,i}^{macro}) + BIGM \cdot (1 - \mathbf{y}_{n,i,k}^{apx}) \end{array} \right) \quad \forall n \in \bar{N}, \forall i \in \bar{I}, \forall k \in \bar{K} \quad (\text{A5})$$

$$\mathbf{akv}_{n,i} \geq \left(\begin{array}{l} LA_{n,i,k}^Y + LA_{n,i,k}^{grad} \cdot (\mathbf{a}_n^{reg} + \mathbf{a}_n^{tmp}) - \\ BIGM \cdot (1 - \mathbf{y}_{n,i}^{macro}) - BIGM \cdot (1 - \mathbf{y}_{n,i,k}^{apx}) \end{array} \right) \quad \forall n \in \bar{N}, \forall i \in \bar{I}, \forall k \in \bar{K} \quad (\text{A6})$$

$$\mathbf{akv}_{n,i} \leq \mathbf{y}_{n,i}^{macro} \quad \forall n \in \bar{N}, \forall i \in \bar{I} \quad (\text{A7})$$

The constraint sets (A2) and (A3) ensure that the $\mathbf{y}_{n,i,k}^{apx}$ variables are set to represent the proper interval. Constraint set (A4) ensures that exactly one interval is selected. In this case $BIGM$ should be selected as the upper bound on the total number of workers. Constraint sets (A5) and (A6) set $\mathbf{akv}_{n,i}$ to its correct value. In this case $BIGM$ should be selected as the upper bound on the total number of workers. The constraints (A7) enforce that the $\mathbf{akv}_{n,i}$ variable can only take a value greater than zero if the corresponding production capacity level is used.

To merge the linear approximation of the working time account with our aggregate planning model, equation (34) must be modified as follows:

$$\mathbf{b}_n = \mathbf{b}_{n-1} + \sum_{m=1}^M \mathbf{z}_{n,m}^r - (1 - A_n^{Abs}) \cdot V_n \cdot \left(1 + \sum_{i=1}^I \mathbf{akv}_{n,i} \right) \quad \forall n \in \bar{N} \quad (\text{A8})$$